3С

УДК 004.93

С.А. Субботин, К.Ю. Бойченко
Запорожский национальный технический университет, Украина
subbotin@zntu.edu.ua

Оценка моделей представления данных в системах обнаружения и распознавания объектов

В статье предложена классификация моделей представления данных в системах обнаружения и распознавания объектов визуальной сцены для решения практических задач. Впервые предложен комплекс критериев для оценки моделей представления данных. Приведены области применения рассмотренных методов.

Введение

Визуальное машинное наблюдение широко применяется в различных областях уже довольно давно. Основным компонентом таких систем наблюдения являются системы распознавания образов, позволяющие интерпретировать визуальные сцены.

В задачах распознавания образов, решаемых на ЭВМ, одной из главных составляющих, влияющих как на качество работы системы распознавания, так и на скорость обработки, является способ представления данных, поступающих непосредственно на вход систем, осуществляющих локализацию и классификацию.

Целью данной статьи является анализ подходов к представлению исходных данных для локализации и распознавания объектов на изображении, что позволит более эффективно выбирать методы обработки и способы представления данных для решения конкретных задач.

Постановка задачи

Изображение сцены и искомого объекта представим в виде двумерных массивов цветовых интенсивностей в каждой дискретной точке. Обозначим эти множества точек S и I соответственно.

Необходимо найти такое представление R множества I, которое позволит эффективно его использовать для широкого круга задач выделения и распознавания объектов.

Также необходимо разработать критерии для оценки применимости моделей представления данных для решения определенных задач.

Методы интерпретации фрагментов визуальной сцены

Рассмотрим задачу интерпретации фрагментов визуальной сцены на примере анализа изображений дорожной сцены с целью обнаружения и распознавания автотранспортных средств. Этот класс распознающих систем использует многие известные на сегодняшний момент методы обнаружения и распознавания объектов.

С целью сокращения размерности поискового пространства для отыскания движущихся объектов на изображении сцены большого размера часто применяется предварительный анализ движения, который осуществляется на основе последовательности кадров, содержащих изображение сцены при неподвижной камере. Один из таких методов описан в [1]. Без привязки к поиску конкретных объектов описанный метод обнаруживает любое движение в пределах изображения сцены.

Далее возможно работать с оригинальным изображением сцены (в том числе, применяя при этом фильтры, улучшающие качество изображения, или уменьшая его масштаб) либо с преобразованным изображением с целью улучшения качества обнаружения и распознавания объектов.

Методы, описанные в [2], [3], позволяют отыскивать как движущиеся, так и неподвижные объекты, используя оригинальное изображение сцены и ряд шаблонов для поиска. В [2] – шаблонов целого автотранспортного средства, снятого с определенного (рассматриваемого) ракурса, а в [3] – семантических фрагментов изображения целого объекта, после обнаружения которых делается вывод как о положении целого искомого объекта, так и о точности обнаружения его фрагментов на основании такого решения.

Существует большое количество различных преобразований изображения, позволяющих облегчить задачу обнаружения и распознавания, повысить их точность, а также выделить инвариантные признаки на основании этих преобразований. Как правило, преобразования связаны с отысканием градиента интенсивности цвета, что сразу делает признаки, основанные на использовании результатов такого преобразования, инвариантными к цвету искомого объекта (что особенно интересно, например, при работе с автотранспортными средствами). Примерами преобразований исходного изображения могут служить сенсус-преобразование [4], карты ребер [5], [6], полученные, например, после применения метода преобразования Собеля, или фильтры Габора [7].

Ряд других методов использует априорную информацию об объекте поиска. Использование кумулятивных гистограмм для обнаружения объектов с известными значениями параметров показывает неплохие результаты при применении в локальной области. Так в [6-8] данный подход был применен для обнаружения автомобилей на видах спереди и сзади, где их форма хорошо вписывается в прямоугольные рамки и имеет сосредоточения ребер, например, горизонтальных в средней части объекта (на границе излома кузова автомобиля).

В [9], [10] описаны методы, содержащие совмещенные процедуры обнаружения и распознавания объектов под любыми углами обзора на основе сопоставления с их трехмерными моделями.

В ряде случаев непосредственно для распознавания наряду с другими, описанными выше методами, применяются нейронные сети [9-12], для обучения которых применялись исходные или преобразованные изображения искомых объектов.

Классификация способов представления данных об объекте поиска

В [13] была сделана попытка классификации методов сегментации изображений. Для нас же представляется целесообразным классифицировать методы обнаружения и распознавания объектов для их применения для решения конкретных задач.

Применяемые в рассмотренных методах обнаружения и распознавания объектов изображения визуальных сцен способы представления данных об объекте поиска различны и зависят от решаемых задач. Условно эти способы можно разделить на несколько основных классов.

На рис. 1 представлена общая схема классификации способов представления изображения для поиска в пределах сцены.

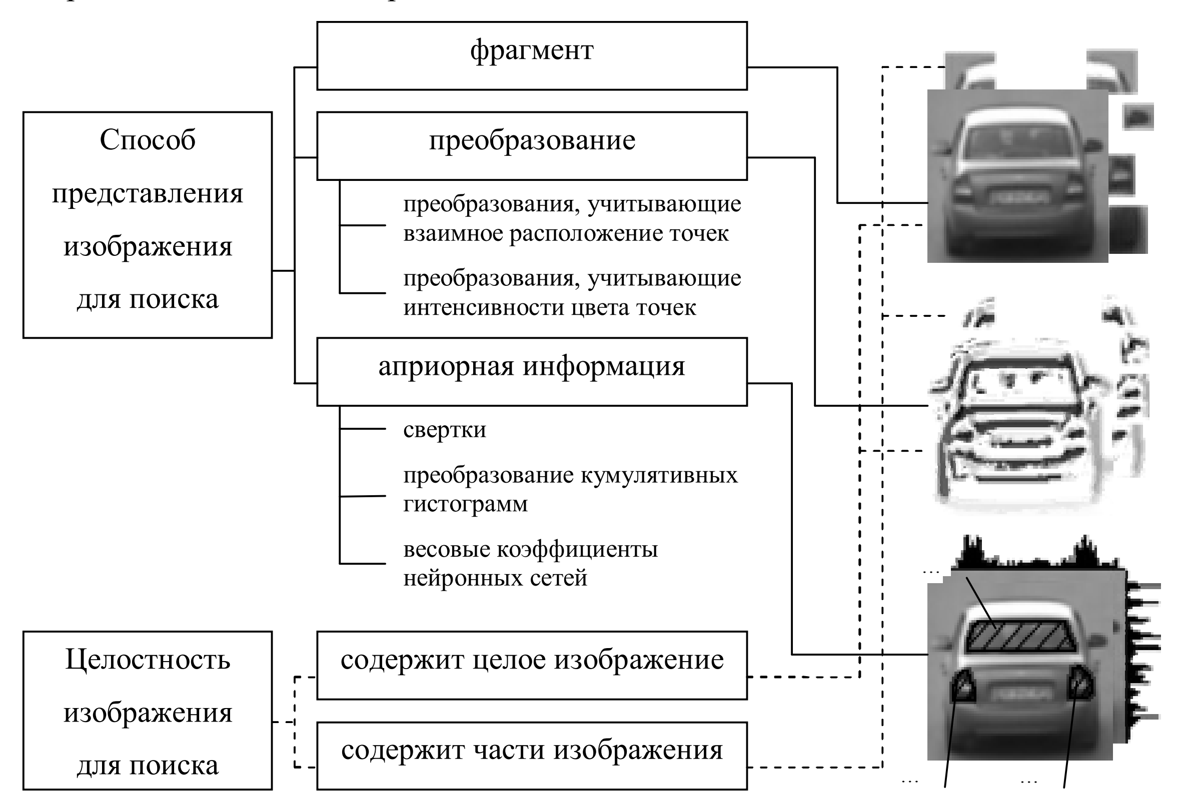


Рисунок 1 – Классификация способов представления изображения для систем оптического распознавания

Самый простой способ представления изображений для поиска – в виде массивов интенсивностей точек, представляющих собой целое искомое изображение

$$R = \{I\}$$

или его фрагменты, после обнаружения которых необходимо провести дополнительные вычисления, сделать вывод на основании взаимного расположения фрагментов о положении искомого объекта:

$$R = \left\{P_k \mid P_k \subseteq I;\ \forall(x_1, y_1) \in P_k, \exists(x_2, y_2) \in P_k, |x_1 - x_2| + |y_1 - y_2| = 1\right\},$$

где P_k – множество дискретных точек; k – индекс; x_1, y_1, x_2, y_2 – натуральные числа.

Преобразования исходного изображения весьма разнообразны и используют информацию об интенсивности и взаимном расположении точек изображения для вычисления параметров каждой такой точки:

$$R = \left\{f\left(I^{(i,j)}, P\right) \middle| P \subset I\right\},$$

где f – некоторая функция преобразования; P – множество дискретных точек; i, j – индексы.

Применение фильтров преобразования [4-7] в различных задачах связано с необходимостью повышения точности обнаружения и распознавания либо сокращения

количества хранимых эталонов классов путем инвариантизации к тем или иным преобразованиям исходного изображения.

При этом применение фильтров, как правило, призвано улучшить качество обнаружения и распознавания без потери большей части первоначальной информации об объекте, в то время как применение сверток предполагает использование методов отбора информативных признаков для решения конкретной задачи.

Найденные таким образом свертки часто могут иметь семантическое значение, упрощая тем самым обработку и интерпретацию работы системы распознавания в целом на более высоком уровне:

$$R = \left\{ f\left(P_k\right) \middle| \, P_k \subseteq I \right\},$$

где f – некоторая функция преобразования; P_k – множество дискретных точек; k – индекс.

Преобразование кумулятивных гистограмм яркости приведено как один из примеров несемантического преобразования с потерей большей части первоначальных данных:

$$R = \left\{ f\left(\vec{c}_1, \vec{c}_2\right) \middle| \, \vec{c}_1 = \left\{ \sum_j g(I^{(i,j)}) \right\} \vec{e}_y, \vec{c}_2 = \left\{ \sum_j g(I^{(j,i)}) \right\} \vec{e}_x \right\},$$

где f, g – некоторые функции преобразования; $\vec{c}_1, \vec{c}_2$ – некоторые векторы; $\vec{e}_x, \vec{e}_y$ – единичные векторы осей декартовой системы координат; i, j – индексы могут применяться с целью ускорения работы системы обнаружения объектов либо полностью ее представлять для отдельных частных задач.

Исходные искомые изображения также можно использовать для обучения нейронных сетей с целью их дальнейшего использования в распознавании. При этом исходные изображения преобразуются в набор весовых коэффициентов обучаемой сети.

Оценка эффективности применения моделей представления данных поискового образца для различных задач

Для оценки эффективности применения моделей представления данных для поиска выделим качественные показатели систем обнаружения и распознавания объектов, на которые оказывает непосредственное влияние модель представления исходных данных для поиска.

Представляется целесообразным обобщить и формализовать результаты сравнения рассмотренных способов хранения, а также очертить круг задач, в которых более рационально использование тех или иных способов.

Результаты сравнения моделей представления данных для поиска приведены в табл. 1.

Как видно из табл. 1, системы, использующие для сравнения изображение объекта или его частей с целью обнаружения и распознавания, являются весьма универсальными, поскольку могут содержать практически любые изображения объектов для поиска. Однако объем хранимых для этого исходных данных велик, поскольку необходимо хранить информацию о цветовой интенсивности в каждой точке каждого искомого изображения объекта.

Таблица 1 – Оценка параметров моделей к представлению данных в системах обнаружения и распознавания объектов

Критерии сравнения моделей представления данных поиска	Фрагмент	Преобразо-вание	Использование априорной информации об объекте поиска	
			Свертки, кумуля-тивные гистограммы и т.д.	Нейронные сети
Универсальность системы.	Высокая	Высокая	Низкая	Низкая
Объем хранимых для поиска данных: – на 1 экземпляр, – количество экземпляров.	Высокий Высокое	Высокий Среднее	Низкий Среднее	Низкий Низкое
Скорость работы системы: – в режиме обучения, – в режиме обнаружения и распознавания.	Низкая Средняя	Низкая Средняя	Низкая Низкая	Высокая Низкая
Трудоемкость создания и обслуживания системы: – на этапе создания, – на этапе подготовки к работе.	Низкая Низкая	Средняя Низкая	Высокая Средняя	Высокая Средняя
Точность обнаружения и распознавания объекта: – максимальная точность при заданном объеме входных данных, – возможность оценки точности.	Средняя Средняя	Высокая Средняя	Высокая Средняя	Средняя Низкая

Исходя из этого можно сделать вывод, что применение такого подхода рационально при поиске определенного класса объектов, не отличающихся особым разнообразием на том или ином уровне дискретизации их изображения.

Затраты на создание такой системы низки, поскольку связаны с созданием системы сравнения изображений. Обучение этих систем сводится к получению набора изображений искомых объектов.

Точность обнаружения и распознавания объектов при заданном объеме входных данных – средняя и обуславливается низкой инвариантностью к различным внешним условиям сцены и положению искомого объекта.

Применение систем, основанных на отыскании исходного изображения объекта, возможно при необходимости минимизировать затраты на создание системы при условии не слишком большого многообразия возможных необходимых для отыскания вариантов изображений объекта.

Преобразование исходного изображения кроме повышения точности обнаружения и распознавания зачастую делает его инвариантным условиям среды или частным случаям изображения искомого объекта, что снижает необходимое для хранения

количество экземпляров преобразованного изображения объекта для достижения заданного качества обнаружения и распознавания. Зачастую в рассмотренных примерах такие преобразованные изображения инвариантны к освещенности сцены и цвету самого объекта, так как используют информацию о градиенте цвета.

Таким образом, использование преобразований связано с необходимостью обнаружения объектов при различных внешних условиях, а также большего числа различных вариантов изображения искомого объекта по сравнению с предыдущим подходом.

Если рассмотренные выше подходы, будучи реализованными для определенного класса задач, могут применяться и в других сферах, то методы, использующие априорную информацию об искомом объекте, разрабатываются для обнаружения и распознавания лишь определенного класса объектов.

Использование априорной информации об искомом объекте позволяет сокращать общее количество признаков, применять обобщение и редукцию, тем самым снижая объем хранимых об объекте данных и времени на обработку при минимальных потерях в качестве работы системы, но требует значительно больших затрат как на этапе создания системы (выделение соответствующих признаков), так и на этапе подготовки к работе (подготовка эталонных изображений в соответствии с выбранными наборами признаков).

Необходимо также отдельно выделить применение для обнаружения и распознавания объектов системы, использующие обученные на изображениях искомых объектов нейронные и нейро-нечеткие сети [12]. Построение таких сетей, как и используемые для их обучения признаки, непосредственно зависят от искомого объекта и строятся на основе априорной информации о нем.

Выводы

С целью решения актуальной задачи распознавания образов предложена классификация способов представления данных в системах обнаружения и распознавания для использования при решении практических задач.

Научная новизна результатов работы заключается в том, что:

– предложен новый метод классификации моделей представления данных в системах обнаружения и распознавания объектов на изображении визуальной сцены для решения практических задач;

– впервые предложены критерии для оценки параметров различных подходов к представлению данных в системах обнаружения и распознавания объектов.

Практической ценностью работы является предложение использования для решения задач обнаружения и распознавания объектов методов, методов и моделей представления данных, наиболее подходящих для данного типа задач.

Работа выполнена как часть НИР кафедры программных средств Запорожского национального технического университета «Научно-методические основы и математическое обеспечение для автоматизации и моделирования процессов управления и поддержки принятия решений на основе процедур распознавания и эволюционной оптимизации в нейросетевом и нечеткологическом базисах» (№ гос. регистрации 0106U008621), а также НИР ООО «МПА Групп» «Разработка математического и информационного обеспечения интеллектуальной системы визуального контроля транспортных средств» (№ гос. регистрации 0106U012013) и «Разработка методов и программных средств на основе обучения, распознавания, оптимизации и адаптации для принятия решений в автоматизированных системах управления транспортными средствами» (№ гос. регистрации 0107U0006781).

Литература

1. Zhou J. Moving vehicle detection for automatic traffic monitoring / J. Zhou, D. Gao, D. Zhang // IEEE Transactions on Vehicular Technology. – 2007. – Vol. 56, № 1. – P. 51-59.
2. Bai L. Computer vision techniques for traffic flow computation / L. Bai, W. Tompkinson, Y. Wang // Pattern Analysis & Applications. – 2004. – Vol. 7, № 4. – P. 365-372.
3. Leibe B. Learning semantic object parts for object categorization / B. Leibe, A. Ettlin, B. Schiele // Image and Vision Computing. – 2008. – Vol. 26, № 1. – P. 15-26.
4. Zabih R. Non-parametric Local Transforms for Computing Visual Correspondence / R. Zabih, J. Woodfill // Proceedings of the Third European Conference on Computer Vision. – 1994. – Vol. 2. – P. 151-158.
5. Betke M. Real-time multiple vehicle detection and tracking from a moving vehicle / M. Betke, E. Haritaoglu, L.S. Davis // Machine Vision and Applications. – 2000. – Vol. 12, № 2. – P. 69-83.
6. Zehang Sun. Monocular precrash vehicle detection: features and classifiers / Zehang Sun, G. Bebis, R. Miller // IEEE Transactions on Image Processing. – 2006. – Vol. 15, № 7. – P. 2019-2034.
7. Shioyama T. Segmentation and object detection with gabor filters and cumulative histograms / T. Shioyama, H. Wu, Sh. Mitani // IEEE 10th International Conference on Image Analysis and Processing. – 1999. – P. 412-417.
8. Vehicle detection with projection histogram and type recognition using hybrid neural networks / [Yi. Liu, Zb. You, L. Cao, X. Jiang] // IEEE International Conference on Networking, Sensing & Control. – 2004. – P. 393-398.
9. Vehicle recognition using boosting neural network classifiers / L. Xia // IEEE 6th World Congress on Intelligent Control and Automation. – 2006. – P. 9641-9644.
10. Wu W.A method of vehicle classification using models and neural networks / W. Wu, Zh. QiSen, W. Mingjun // IEEE Vehicular Technology Conference. – 2001. – Vol. 4. – P. 3022-3026.
11. Mantri S. A neural network based vehicle detection and tracking system / S. Mantri, D. Bullock // IEEE Twenty-Seventh Southeastern Symposium on System Theory. – 1995. – P. 279-283.
12. Субботин С.А. Синтез нейро-нечетких моделей для выделения и распознавания объектов на сложном фоне по двумерному изображению / Сергей Иванович Субботин // Комп'ютерне моделювання та інтелектуальні системи : [зб. наук. пр. / ред. Пізи Д.М., Субботіна С.О.]. – Запоріжжя : ЗНТУ, 2007. – С. 68-91.
13. Zhang Y.J. A review of recent evaluation methods for image segmentation / Y.J. Zhang // IEEE International Symposium on Signal Processing and its Applications. – 2001. – P. 148-151.

С.О. Субботін, К.Ю. Бойченко

Оцінка моделей подання даних в системах знаходження та розпізнавання об'єктів

У статті запропоновано класифікацію моделей подання даних в системах знаходження та розпізнавання об'єктів візуальних сцен для вирішення практичних задач. Вперше запропоновано комплекс критеріїв для оцінки моделей подання даних. Наведено галузі застосування оглянутих методів.

S.A. Subbotin, K.Yu. Boichenko

Data Representation Models Evaluation in Object Detection and Recognition System

The data representation models classification in visual scene object detection and recognition systems is introduced. The criteria complex for data representation model evaluation is developed. The application domains for reviewed methods are adduced.

Статья поступила в редакцию 22.10.2008.